Zhe Zhang

Systems Research Lab, EB II 3226,
Department of Computer Science
NC State University, Raleigh, NC 27695
http://www4.ncsu.edu/~zzhang3/

3100 Kings Ct. Apt. F
Raleigh, NC 27606
(919) 412-8828
zzhang3@ncsu.edu

Research Interests

My research mainly focuses on the area of *distributed computing systems*, with specific interests in *distributed file systems* and *caching/prefetching techniques*.

Education

NC State University, Raleigh, NC — 2009
Ph.D. in Computer Science and Operations Research (GPA: 4.0/4.0)
Advisor: Dr. Xiaosong Ma
Dissertation Topic:
Adding Coordination to HEC Storage Stacks for Efficiency and Reliability

NC State University, Raleigh, NC — 2006
Master of Science in Operations Research (GPA: 4.0/4.0)
Advisor: Dr. David F. McAllister
Thesis: *Application of Linear Optimization in Stereo Image Rendering*

University of Science and Technology of China, Hefei, China — 2003
Bachelor of Engineering in Computer Science (Avg. Score: 83/100)

Professional Experience

NC State University, Raleigh, NC — Aug. 2004∼present
Research Assistant

Microsoft Research, Cambridge, UK — Jul.∼Sep. 2008
Research Intern

IBM T.J. Watson Research Center, Hawthorne, NY — May∼Aug. 2008
Research Intern

Cisco Systems, San Jose, CA — May∼Aug. 2007
Software Engineering Intern, Service Routing Group

Oak Ridge National Laboratory, Oak Ridge, TN — Jun.∼Aug. 2006
Research Intern, Distributed Systems Research Group

Publications

Refereed Conferences

1. **[Middleware '09]** *"An Empirical Study of High Availability in Stream Processing Systems"* by Yu Gu, *Zhe Zhang (First 2 authors in alphabetical order)*, Fan Ye, Hao Yang, Minkyong Kim, Hui Lei and Zhen Liu. To appear in the Proceedings of *ACM/IFIP/USENIX International Middleware Conference (Middleware), Industrial Track*, Urbana Champaign, Illinois, Dec. 2009.
2. **[JGC '09]** *"Improving Data Availability for Better Access Performance: A Study on Caching Scientific Data on Distributed Desktop Workstations"* by Xiaosong Ma, Sudharshan Vazhkudai, and *Zhe Zhang*. Published in the *Journal of Grid Computing (JGC), Special Issue on Volunteer Computing and Desktop Grids* , 2009.
3. **[ISC '09]** *"Improving the Availability of Supercomputer Job Input Data Using Temporal Replication"* by Chao Wang, *Zhe Zhang*, Sudharshan Vazhkudai, Xiaosong Ma

and Frank Mueller. Published in the Proceedings of *International Supercomputing Conference (ISC)*, Hamburg, Germany, Jun. 2009.

4. **[EuroSys '09]** *"Memory Resource Allocation for File System Prefetching – From a Supply Chain Management Perspective"* by *Zhe Zhang*, Amit Kulkarni, Xiaosong Ma and Yuanyuan Zhou. Published in the Proceedings of *European Conference on Computer Systems (EuroSys)*, Nuremberg, Germany, Apr. 2009 (**17% acceptance rate**).
5. **[ICPP '08]** *"On-the-fly Recovery of Job Input Data in Supercomputers"* by Chao Wang, *Zhe Zhang*, Sudharshan Vazhkudai, Xiaosong Ma and Frank Mueller. Published in the Proceedings of *International Conference on Parallel Processing (ICPP)*, Portland, Oregon, Sep. 2008
6. **[ICDCS '08]** *"PFC: Transparent Optimization of Existing Prefetching Strategies for Multi-level Storage Systems"* by *Zhe Zhang*, Kyuhyung Lee, Xiaosong Ma and Yuanyuan Zhou. Published in the Proceedings of *IEEE International Conference on Distributed Computing Systems (ICDCS)*, Beijing, China, Jun. 2008 (**16% acceptance rate**).
7. **[SC '07]** *"Optimizing Center Performance through Coordinated Data Staging, Scheduling and Recovery"* by *Zhe Zhang*, Chao Wang, Sudharshan Vazhkudai, Xiaosong Ma, Gregory G. Pike, John W. Cobb and Frank Mueller. Published in the Proceedings of *ACM/IEEE Supercomputing (SC)*, Reno, NV, Nov. 2007 (**20% acceptance rate**).
8. **[EI '06]** *"A Uniform Metric for Anaglyph Calculation"* by *Zhe Zhang* and David F. McAllister. Published in the Proceedings of *IS&T/SPIE Electronic Imaging (EI)*, San Jose, CA, Jan. 2006.

PATENT

[HA-patent '09] *"Method and System for High Availability in Distributed Messaging and Stream Processing Systems"* by Yu Gu, Minkyong Kim, Zhen Liu, Hao Yang, Fan Ye and *Zhe Zhang* (in alphabetical order). US patent pending.

PROJECTS

MapReduce in opportunistic environments: Collaboratively researched a hybrid architecture and multiple novel techniques to enable Hadoop to run MapReduce applications on resource-scavenging grids with high reliability. Developed automatic tests to evaluate the performance of the system on large platforms.
Work conducted at NCSU supervised by Dr. Xiaosong Ma(NCSU/ORNL).

Rate-aware File System Prefetching: Innovatively observed the similarity between file system prefetching and supply chain management (SCM), and performed mapping of concepts between the two areas. Based on that theoretical analysis, proposed and designed novel mechanisms to measure data access rates and allocate file system memory space among prefetching streams accordingly. The prefetching algorithms are implemented in the Linux 2.6.18 kernel and are able to improve the throughput of mixed workload by up to 33%.
Work conducted at NCSU supervised by Dr. Xiaosong Ma(NCSU/ORNL) and Dr. Yuanyuan Zhou(UIUC), and I was one of the two major developers.
Related publications : [EuroSys '09]

Multi-level Buffer Cache Management: Proposed and contributed significantly to the design of PreFetching Coordinator(PFC), a novel multi-level coordinated prefetching mechanism. Rather than being another new prefetching algorithm, PFC acts as a middleman between two adjacent levels of caching/prefetching and optimizes behaviors of existing algorithms. A C++ based multi-level cache simulator is developed and attached to DiskSim to measure the performance of PFC working with different prefetching algorithms including AMP, SARC, and the algorithm used by Linux 2.6 kernel. Ongoing work aims at implementing PFC into the Lustre parallel file system.
Work conducted at NCSU supervised by Dr. Xiaosong Ma(NCSU/ORNL) and Dr. Yuanyuan Zhou(UIUC), and I was one of the two major developers.

Related publications : [ICDCS '08]

High Availability in Overlay Messaging and Streaming Systems: Proposed and designed a hybrid failure tolerance method for providing high availability to overlay messaging and streaming systems, which combines advantages of existing active standby and passive standby methods and allows users to flexibly choose desirable positions in the overhead-delay tradeoff. Also made major contribution in building a software system for evaluating and analyzing different fault tolerance techniques for overlay systems. The system is developed in Java, has been tested on IBM's System S cluster, and has a prospective deployment on larger platforms of IBM's customers.
Work conducted at IBM T.J. Watson Research Center in the Next Generation Distributed Systems Group led by Dr. Zhen Liu, supervised by Dr. Fan Ye(IBM), Dr. Hao Yang(IBM) and Dr. Minkyong Kim(IBM), and I was one of the three major developers.
Related publication/patent : [Middleware '09, HA-patent '09]

Fault Tolerance in Distributed Storage Systems: Developed a prototype data staging/reconstructing manager based on the Moab job scheduler and the Lustre parallel file system. Our prototype allows HPC users to specify data operations in their PBS scripts, and uses job scheduling information to achieve just-in-time data staging. We also made modifications to the Lustre parallel file system to transparently patch data from remote sources and reconstruct files from partial data loss. With a proposed deployment on the ORNL Jaguar system(No. 7 in the top500 list), the prototype has been tested on multiple clusters at ORNL and NCSU.
Work conducted at NCSU supervised by Dr. Sudharshan Vazhkudai (ORNL), Dr. Xiaosong Ma (NCSU/ORNL), and Dr. Frank Mueller (NCSU), and I was one of the two major developers.
Related publications : [ISC '09, ICPP '08, SC '07]

Distributed Service Routing: Implemented new commands for users to configure and control the Distributed Hash Table(DHT) based service routing system. Improved system fault tolerance and performance by adding backup bootstrap points to accept new nodes joining the network. The system mainframe is in C and the commands are written in Java.
Work conducted at Cisco Systems supervised by Albert Tian(Cisco).

Distributed Storage Scavenging: Designed and implemented an on-line cache management algorithm for *FreeLoader*, a novel distributed storage scavenging file system. FreeLoader aggregates underutilized desktop storage space to provide a shared cache/scratch space for large, immutable data sets. Based on the access patterns of this kind of files, my algorithm assigns priorities to data blocks based on both their access recencies and logical offsets. I also contributed to the implementation of parallel data transmission using TCP/IP sockets.
Work conducted at ORNL supervised by Dr. Sudharshan Vazhkudai(ORNL) and Dr. Xiaosong Ma(NCSU/ORNL)
Related publication : [JGC'09]

Stereo Image Rendering: Designed an MPI/C++ program to implement a stereo image rendering algorithm. Applied depth-first traversing technique to process areas of similar colors in real world images at low cost(workload reduced by 72% to 89% in extensive tests).
Work conducted at NCSU supervised by Dr. David F. McAllister(NCSU).
Related publication : [EI '06]

HONORS & MEMBERSHIPS:

Phi Kappa Phi, Elected as **top 10%** graduate student for 2004∼2007 academic years
Passed Computer Science Ph.D. written and oral preliminary exams in 2008
High Pass in Operations Research Ph.D. qualifying exam in 2005
Outstanding Student Scholarship, Grade 2, from USTC in 2002(**top 12%** students).

Professional Service & Membership

Reviewer, the 2008 IEEE International Conference on Distributed Computing Systems (ICDCS '08)

Reviewer, the 2008 IACC International Conference on Parallel Processing (ICPP '08)

Reviewer, the 2008 ACM/IEEE Supercomputing Conference (SC '08)

Reviewer, the 2008 ACM Statistical and Scientific Database Management Conference (SSDBM '08)

Reviewer, the 2007 IEEE International Conference on Networking, Architecture, and Storage (NAS '07)

Reviewer, the 2006 ACM/IEEE Supercomputing Conference (SC '06)

Student Members: ACM, SIAM since 2004

Teaching

Teaching Assistant, **NC State University** (Aug. 2004∼May 2006)
In the undergraduate courses *Numerical Analysis*, *Computational Theory* and *File Organization*, assisted the instructors by holding office hours, grading and giving lectures.

Skills

Linux, MS-Windows, C, C++, Java, Pascal
Perl, CGI, PHP, JSP, OpenGL, MPI, OpenMP, SQL, JDBC

References

Dr. Xiaosong Ma
Assistant Professor
Department of Computer Science
North Carolina State University
Joint Faculty
Computer Science and Mathematics Division
Oak Ridge National Laboratory
Phone: 919-513-7577
Email: ma@csc.ncsu.edu

Dr. Sudharshan Vazhkudai
Research Staff Member
Computer Science and Mathematics Division
Oak Ridge National Laboratory
Phone: 865-576-5547
Email: vazhkudaiss@ornl.gov

Dr. Frank Mueller
Associate Professor
Department of Computer Science
North Carolina State University
Email: mueller@cs.ncsu.edu

Dr. Fan Ye
Research Staff Member
IBM T.J. Watson Research Center
Phone: 914-784-7299
Email: fanye@us.ibm.com

Dr. Yahya Fathi
Professor
Edward P. Fitts Department of
Industrial & Systems Engineering
Co-Director
Operations Research Program
North Carolina State University
Email: fathi@ncsu.edu